DOCUMENT RESUME

ED 421 626 CE 076 852

AUTHOR	Barnard, Philip; Dixon, Stella
TITLE	Value Added: Beyond A-Levels to Vocational Programmes?
INSTITUTION	Further Education Development Agency, London (England).
ISSN	ISSN-1364-6869
PUB DATE	1998-08-00
NOTE	17p.
AVAILABLE FROM	Information Centre, Further Education Development Agency, Citadel Place, Tinworth Street, London SE11 5EH, United Kingdom (3.50 pounds).
PUB TYPE	Collected Works - Serials (022) -- Reports - General (140)
JOURNAL CIT	FEDA Bulletin; v2 n4 1998
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Accountability; Comparative Analysis; Evaluation Methods; Feasibility Studies; Foreign Countries; *General Education; *Measurement Techniques; Outcomes of Education; Postsecondary Education; *Program Effectiveness; Secondary Education; Technical Institutes; *Vocational Education
IDENTIFIERS	General National Vocational Qualif (England); *Great Britain; *Value Added Model

ABSTRACT

In Great Britain, interest in the measurement of value added was stimulated initially by the requirement that schools and further education (FE) colleges publish league tables of examination results. Value-added data can be aggregated to compare different institutions' performance. Within colleges, value-added data provide a basis for setting targets, monitoring performance, and identifying patterns of success and failure. Measuring value added entails comparing the characteristics and attainments of learners at entry (input data) and their achievements at exit (output data). Where correlation between input scores and output scores is high, value-added data can also be used to motivate and improve the performance of individual students by setting realistic targets and monitoring progress. According to a recent study that measured the performance of nearly 2,000 students in General National Vocational Qualifications (GNVQ) advanced programs at 10 FE colleges, nationwide application of a uniform methodology for the measurement of value added on vocational courses is not presently feasible. The General Certificate of Secondary Education is the best predictor of GNVQ attainment. Colleges should not, however, be discouraged from using local value-added measurements, provided those measurements are based on a sound statistical methodology. (15 references) (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 421 626

FEDA Further Education Development Agency

bulletin

Volume 2 Number 4

ISSN 1364 6869 Copyright © FEDA 1998

Value added: beyond A-levels to vocational programmes?

Philip Barnard and Stella Dixon

August 1998 £3.50

Contents

Summary	1
The meaning of value added	2
Current uses of value added in colleges	5
The feasibility of measuring value added for students on GNVQ programmes	8
The way forward	14

Summary

Value added means different things to different people. For some it concerns the benefits, besides qualifications, which students gain from attending college. Thus greater confidence, maturity and life skills are seen as added value. For other people value added entails the use of statistical techniques for comparing the attainments of students on entry with their achievements at exit. Both uses of the word are, of course, valid. This publication, however, is about the second, technical, use of the term.

This bulletin:

- explores the meanings of value added
- outlines its current use in colleges
- reports on a feasibility study to see if the value added methodology can be applied to GNVQs
- suggests ways forward.

It will be of interest to:

- national agencies
- senior managers
- teachers with an interest in value added.

widening participation

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
S Barnard
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

E 076 852

2

The meaning of value added

Context

Interest in the measurement of value added was stimulated initially by the requirement for schools and colleges to publish league tables of examination results. It is widely recognised that examination attainments alone do not fairly reflect the academic performance of schools and colleges, since institutions which start off with students with higher levels of prior attainment may well be expected to achieve better results. Measurements based on value added allow for fairer comparisons.

These measurements take into account the starting points of pupils and students, and use them to measure whether they have done as well as, worse than, or better than would have been expected.

Value-added data can be aggregated to compare the performances of different institutions. If a clear pattern of relationships is found between starting points and subsequent achievements, the data can also be used to predict the probable performance of individual students. Within colleges value-added data therefore provide a basis for setting targets, monitoring performance, and identifying patterns of success and failure.

Definition and example

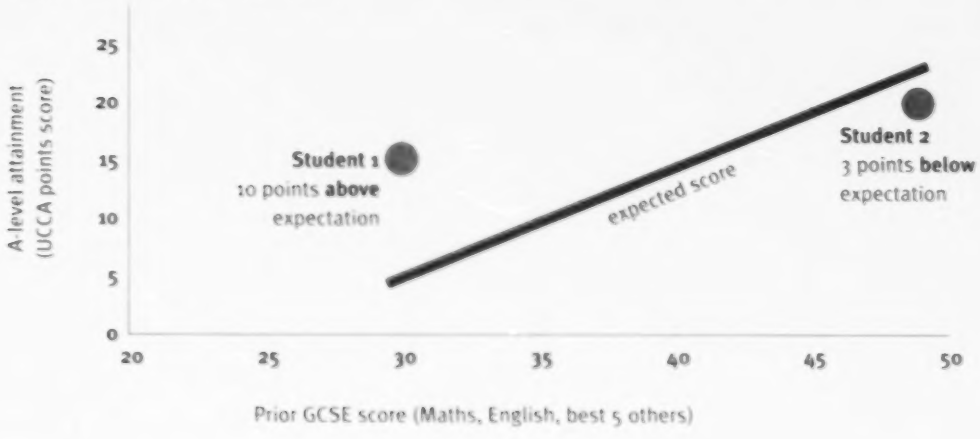
The measurement of value added entails comparison between the characteristics and attainments of learners at entry (the input data) and their achievements at exit (the output data).

This relationship is perhaps most easily understood when it is represented by plotting individual scores on a graph, with input data scaled on the horizontal axis and output data on the vertical axis. Using a statistical technique known as regression analysis, a line can be drawn on the graph to show the general relationship between inputs and outputs.

When value-added findings are being used to provide guidance and set targets for individual students, it may be useful to set out the data in the form of chances charts. These can be used to help students assess their own chances of success. Chances charts show the percentage of students with a given average GCSE score who achieve various A-level grades for a subject. Figure 2 illustrates this for Physics, opposite.

Average GCSE scores commonly account for between 36% and 49% of the variation in GCE A-level scores. A much lower value has been found for the relationship between average GCSE scores and levels of attainment on vocational courses. Value-added data may therefore be used to make comparisons between or set targets for students on vocational courses only with great caution or not at all.

Figure 1 The concept of value added



Thus **student 2** has better A-level results than **student 1** but a less good value-added result.

1993 Unfinished Business. HMSO

Figure 2 ALIS chances charts for GCE A-level Physics

Figure 2a Average GCSE between 4 and 4.75

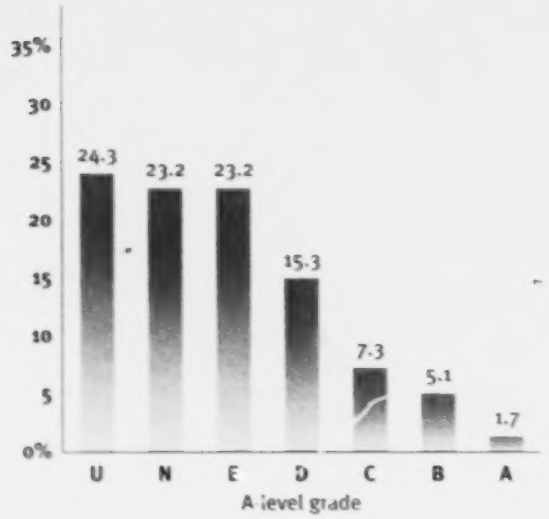


Figure 2b Average GCSE between 4.75 and 5.5

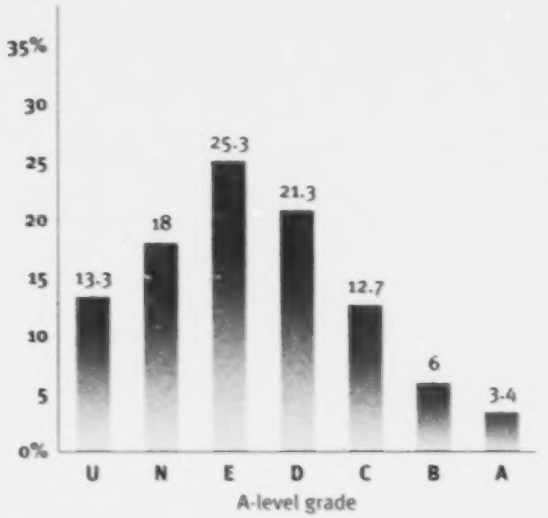


Figure 2c Average GCSE between 5.5 and 6.25

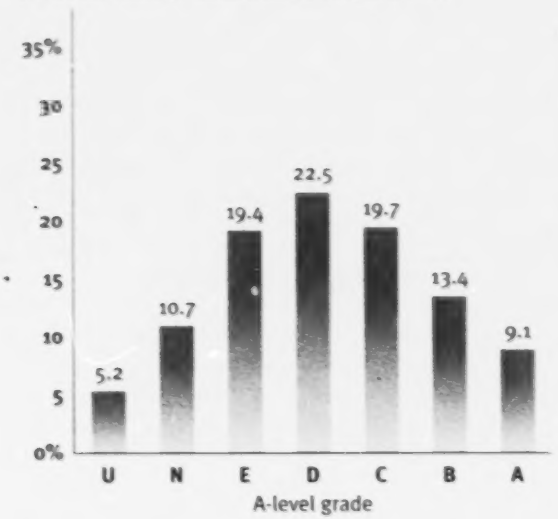
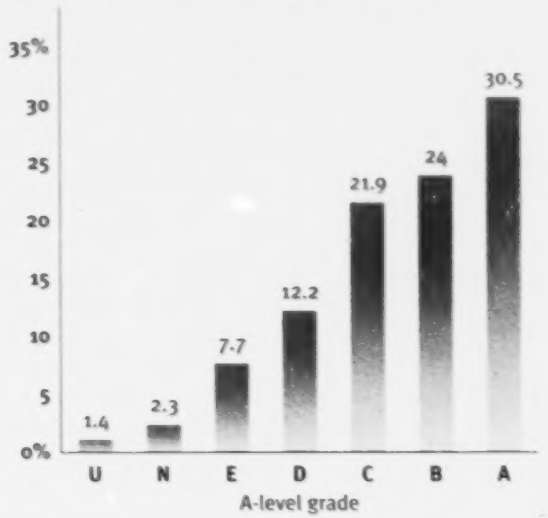


Figure 2d Average GCSE between 6.25 and 7



As is implied by the chances charts, measurements of value added are expressed in terms of probabilities. In statistical terms, the strength of the relationship between inputs and outputs is measured by the 'correlation coefficient' (denoted by the letter 'r'). This measure ranges between $r = 0$, where there is no statistical relationship between inputs and outputs, and $r = 1$ (or inversely $r = -1$), where the relationship of inputs to outputs is completely consistent.

In measurements of the relationship between average GCSE scores and GCE A-level scores, the value of r is commonly between 0.6 and 0.7. The value of r^2 indicates the extent to which the variation in the output score is determined by the input score.

Benefits and limitations

Where there is a high correlation between input scores and output scores, value-added data can be used both to make fair comparisons between the performances of different schools and colleges, and to motivate and improve the performance of individual students by setting realistic targets and monitoring progress.

In many colleges the opportunities for using value-added data are limited. Most college programmes consist of relatively small groups of students following a diversity of courses and options. Findings based on data provided by small groups tend to have poor statistical validity. Roughly speaking, statistically valid findings need to be based on groups of 30 or more students following identical programmes.

It may be difficult to marshal the resources and expertise needed to carry out value-added analysis in colleges. The input data available from older students tend to be more varied – and sometimes less accurate – and the output data may be either difficult to score consistently or be based on a simple differentiation between pass and fail which is not susceptible to statistical analysis. Highly organised data collection and advanced statistical expertise are required to address these problems.

However, staff in a number of colleges have felt that the advantages of setting target grades to motivate and monitor the performance of students outweigh the disadvantages of statistically weak prediction. Provided that the student support system is flexible, the setting of target grades which are based on rather weak statistical evidence may help to provide a structure and focus, and thereby offer a valuable tool for improving performance. It should not be forgotten that the virtue of such a system lies in the good relationship between the tutor and student, not in the value-added methodology.

Evidence from practice in some colleges, and even from a few more sophisticated statistical studies, shows that it is very easy to fall into the error of assuming that students' previous levels of attainment will be reflected in their subsequent performance, even though there is no statistical evidence for this. Comparisons made on the basis of such unsubstantiated assumptions may be just as unfair as comparisons based on raw examination results.

In its work on value added over the past four years, FEDA has sought to identify a methodology which combines practicability with sufficient statistical rigour. The project research has made it increasingly clear that there is a limited range of situations in which it is proper to apply value-added methods. In such situations value added is a powerful tool. In other situations the identification of strengths and weaknesses and the improvement of performance may better be sought by other methods of quality assurance.

Current uses of value added in colleges

Context

Over the past few years the emphasis of interest in value added has ranged from league tables comparing the value-added performance of different schools and colleges to practical applications of value-added data to help students and staff to improve retention and achievement. Interest in the latter has been reinforced by the incentive in the funding methodology to improve retention and achievement, and by the interest of the inspectorate in well designed systems for quality improvement. The extent of this range may be seen in terms of the contrast between 'proving' value for money to an external audience and 'improving' performance within the college.

Value-added findings can be used both with students and with staff, and may point to ways of improving the delivery of a wide range of programmes as well as the teaching of individual subjects. The following descriptions of current practice are based on the use in colleges of GCSE results as value-added predictors for students who are following courses leading to GCE A-levels.

Using value added with students

The most effective way of using value-added data with individual students is in the context of a well planned student tutorial system. Forms of support which make use of value-added data almost inevitably require one-to-one contact between tutor and student. When introducing the use of value-added data within an existing tutorial system, it may be necessary both to train staff and to reallocate responsibilities and resources.

For example, a sixth-form college in the south of England has over the past few years set up a tutorial and teaching system which makes extensive use of value-added data. In developing this system it has gone through the following processes:

- a review of systems currently in use in other colleges for measuring and applying value added
- an identification of the data needed by the college as a basis for the provision of support based on value added
- a complete overhaul of the college's tutorial system, leading to the allocation of tutorial responsibilities to a smaller number of specially designated tutors led by a cross-college senior tutor

- the provision of value-added data in a format which can be readily understood by staff some of whom may have a limited understanding of statistical techniques
- the development of user-friendly spreadsheet software which enables subject tutors easily to monitor the value-added performance of students as part of the process of marking assignments.

The availability of support from the college's management information system (MIS) and the development of the attitudes and skills of individual subject teachers have both been key factors in this development.

With individual students value-added data may be used to:

Set targets on entry

Appropriate value-added norms must be identified as a basis for target-setting. National norms for value added are available in DfEE publications (1995, 1997). These take very limited account of factors affecting specific subjects and specific colleges. Norms for individual subjects based on a national sample are provided by the A-Level Information System (ALIS). Data collected in an individual college over a number of years may be robust enough to allow the adoption of valid college norms.

In discussing target minimum grades with a student on entry, it may be necessary to distinguish clearly between a target based on statistical prediction and a target based on the student's personal circumstances and aspirations. In some cases, the use of chances charts may help a student to appreciate this distinction. In all cases, it should be made clear that the normatively based target grade should be viewed as a minimum target grade.

Review performance against targets

Opinion varies between different colleges on how frequently tutors should review with students their performance against target grades. In some colleges the review is undertaken every half term, in others less frequently. Most colleges find that their students need to go through a settling down period, and that judgements about performance in the early stages may not be helpful. One way of addressing this situation is to avoid awarding examination-based grades during the first period of review, and to assess performance more generally in terms of how well the student is coping with the course.

Identify support needs

Value-added targets can be used effectively in a tutorial system only if support can be offered to students who are failing to achieve their target grades. If a student's underperformance is due to a lack of general learning skills, support may need to be given in workshops for the development of language, numeracy, study skills or essay-writing. Where the need for support is subject-specific, it may be necessary to provide subject workshops in addition to the normal timetabled provision.

Underperformance may also be due to personal circumstances or motivational factors, in which case referral to student-care support or to more senior tutorial staff may be appropriate.

Agree individual action plans

Once the reasons for underperformance have been identified by the tutor and student, a plan of action should be agreed. This may involve the tutor taking action to sanction a referral; the student and/or tutor following the agreed action; and the monitoring of the effectiveness of the action in further tutorial contact.

Monitor effectiveness

The use of targets based on value-added data entails the deployment of significant resources, including information technology (IT) support, tutorial provision, and workshops. There is evidence of a clearer sense of purpose and enhanced motivation in colleges which have established a target-based tutorial support system. However, it is sometimes more difficult to obtain hard evidence of such systems leading to improvements in academic achievements.

Using value-added data at subject level

In providing target grades for individual students, value-added norms are used as predictors of expected levels of achievement. The value-added data for a group of students taught in the same college can also be used retrospectively to identify patterns of achievement.

In a tertiary college in the West Midlands, for example, staff teaching modern languages found, when analysing the results of of their students who had failed to achieve the GCE A-level grades predicted by value-added norms, that the majority of their underperforming students had previously attended one specific partner secondary school.

On reflection, it appeared to college staff that the pupils from this school were not being adequately prepared to cope with the demands for grammatical accuracy made in the A-level courses.

The college was able to take action by offering support to the modern language teachers in the school. Lecturers undertook some teaching with the older pupils in the school; teachers came in to participate in teaching in the college; and eventually arrangements were set up for pupils to make use of the college's language learning facilities.

Not all findings of patterns of strengths and weaknesses can lead to such thorough-going action. One college found that A-level physics students with high average GCSE scores but with C grades in GCSE double science did less well than students with the same average score, but with higher grades in science. A college which participated in a FEDA value-added project found that students' failure to complete an NVQ Level 2 Hairdressing course almost exactly matched their failure to achieve above Level One in a basic skills screening test administered at the start of their course.

In each of these cases, the identification of a pattern in value-added data enabled staff to form a hypothesis to account for underperformance, and provided a basis for taking action within the college to improve achievement. The stages involved are:

- recognition of patterns
- generation of hypotheses to account for these patterns
- decision to take specific action based on hypotheses
- testing the hypotheses by monitoring the effectiveness of the action.

Using value-added data at college level

Patterns of over- or under- achievement may also be identified when results for a group of subjects are analysed. A sixth form college involved in a FEDA project discovered that across a range of A-level subjects it added value more consistently among stronger and weaker students than to those who came in with GCSE scores in the middle range.

Staff then recognised that college-wide systems already provided support for students with poor GCSE scores on entry and a mentor system to stretch very able students, but there was not systematic support for students in the middle range.

To address this situation, the college considered devising staff development programmes which focused on strategies to engage the average student.

Another college found that students who were undertaking paid work outside the college for more than 10 hours per week tended to underperform. This was a finding which staff could use to give students guidance when they were being interviewed for admission and during the induction programme.

A number of colleges report that students who have been intensively coached in their secondary schools to achieve high GCSE grades tend to find it difficult to develop the more independent study skills required for success at A-level. Here again, the identification of a pattern and the generation of a hypothesis to explain it may provide the basis for effective remedial action.

Using value added in staff management

With the increasing emphasis on improving retention and achievement, there is increasing pressure to ensure that staff are making a contribution to adding value. Implementing the measurement of value added is one way in which colleges may seek to identify strengths and weaknesses in teaching and learning. This places an onus on staff whose students are underachieving. Value-added measurement itself cannot offer any explanation for such underachievement; it may be due to the characteristics of the learners or the input of the teachers or a combination of both.

It is reasonable to expect that teaching staff should be able to offer a plausible explanation for such underachievement, to propose action and to set targets for improvement. In some colleges, reporting by staff to senior management on the value-added performance of their students is already well established as part of an ongoing quality assurance process. It is important that a self-critical, improving culture is established to ensure the successful implementation of such measures.

Key messages

Value-added data can be used to set targets and help individual students to improve their performance only in the context of a well developed student support system. It requires tutorial staff who are able both to offer the interpersonal skills needed to motivate students by negotiating realistic targets and agreeing effective courses of action, and to cope with elementary statistical concepts. Tutorial support needs to be backed up by other forms of support to address personal needs, to provide basic skills support, and to offer subject-specific tuition.

The use of value-added analysis to identify patterns of strength and weakness normally requires that explanations should be put forward by staff who have first-hand knowledge of the characteristics of their students. Such explanations and proposals for appropriate action may therefore best be incorporated in the self-assessment activities of course teams. In many cases, appropriate action will require support at a more senior management level.

The provision and use of value-added data, whether externally or internally derived, is most efficient if it can be supported from within the college's management information system. Like value-added measurements, college management information systems were initially used almost exclusively to 'prove' the college's performance. With experience, some college information systems managers are beginning to cope more easily with the demand for data, and are developing support for systems which can be used to 'improve' the college's performance. Value-added data can be used most effectively in the context of an information-based college quality management system.

The guidance offered here applies to situations in which reliable measurements of value added within a college are already available, either from data collected and analysed by college staff or through subscription to an external service, such as ALIS. As yet, reliable ways of measuring value added in colleges have been established only for full-time students on GCE A-level programmes.

The feasibility of measuring value added for students on GNVQ programmes

Can value added be measured on vocational programmes?

Summary

From 1994 to 1996 FEDA measured the performance of nearly 2,000 students on GNVQ Advanced programmes at 10 colleges. In addition to GCSE scores, the study included data based on psychometric tests, levels of basic skills, socio-economic factors, and attitudinal measures. The students' achievements on their GNVQ programmes were measured in terms of final grades awarded, and information on additional qualifications and partial completion was included where available.

The study aimed to find out whether value added on vocational programmes could be measured in a way which was statistically valid and did not make disproportionate demands on resources.

Analysis of the data obtained confirmed earlier findings that GCSE scores are poor predictors of performance on vocational programmes. In this study the correlation was just under $r = 0.3$, thus accounting for only about 8–9% of the variation in GNVQ grade, whereas for students on GCE A-level courses their previous GCSE score would typically account for up to 50% of the variation in their grade.

The inclusion of additional factors as input data did not significantly improve the rather poor prediction of outcomes provided by GCSE scores. Their effects appeared to be slight, and the interpretation of these effects was problematic.

The conclusion to be drawn from the experience of this study is that it is not at present feasible to apply nationwide a single uniform methodology for the measurement of value added on vocational courses.

Context

Since the early 1990s the validity of value-added measurements had been demonstrated only for students progressing from GCSEs to GCE A-levels, who constitute a small minority of students in most colleges.

The first attempts to apply the established value-added methodology to vocational courses were disappointing: GCSE scores showed a poor correlation with students' performance in the nearest vocational equivalents of GCE A-level (see Audit Commission/HMI, 1992). Further investigation was required to find out whether alternative or additional input data could be used as input measures to provide an acceptably high level of prediction for use on vocational courses.

The design of the FEU/FEDA study

Evidence from research studies indicated that the best predictor of educational performance at this stage would incorporate a cognitive element providing a measure of mental ability and/or previous academic attainment. The only practicable alternatives to GCSE scores were psychometric test scores and measurements of performance in basic skills tests. Both of these appeared to overcome the most plausible objection to the use of measures based on GCSE – that GCSE was too academic to be used as a predictor on vocational courses.

Value-added research had also shown that predictions based on cognitive measures could be improved by taking into consideration other factors. For example, the prediction of GCE A-level grades from GCSE scores could be improved by making adjustments based on the student's gender and the academic attainments of the student's parents. The additional factors which appeared likely to be associated with differences in achievement on vocational courses were:

- gender and age
- ethnicity and language spoken at home
- socio-economic background
- non-academic commitments, such as caring for children and part-time employment
- attitudes, motivation, and academic and vocational aspirations.

Following a pilot project in 1992–3 and a nationwide survey of college use of value-added data in the Spring of 1994, FEDA's predecessor FEU investigated the feasibility of measuring value added for students on selected GNVQ Advanced programmes starting in September 1994. GNVQ Advanced programmes were chosen because, as the vocational equivalents of A-levels, they both constituted a key element in the further education curriculum and were sufficiently similar to academic A-levels to offer a reasonable prospect of successfully adapting the existing value-added methodology.

The survey had shown that the GNVQ Advanced programmes on which students were enrolled in sufficient numbers to provide useable data for a value-added study were: Art and Design, Business, Health and Social Care, and Leisure and Tourism. Ten colleges were selected to participate in the study. They provided a sample of nearly 2,000 students, and were broadly representative of the FE sector as a whole. Table 1 shows the composition of the sample by college and GNVQ programme.

Table 1 Composition of GNVQ Advanced sample

College	A&D	BUS	HSC	L&T
A	20	26	17	33
B	38	182	89	86
C	51	27	32	69
D	5	48	35	33
E	16	55	22	55
F	–	64	33	13
G	–	71	44	16
H	27	50	28	43
I	31	54	69	40
J	28	240	89	73
Total	216	817	458	461
			Grand total	**1,952**

The range of input data

Psychometric tests

The initial assumption of the researchers had been that psychometric tests of practical abilities would offer the best prediction of students' performance on vocational courses. Expert advice was sought from Professor David Bartram of the University of Hull. The advice he provided was based on an analysis of the formal specifications for the GNVQ programmes.

Professor Bartram advised that tests of general mental abilities appeared to be the most likely predictors of success. In addition, he suggested that spatial abilities might be associated with success on the GNVQ Art and Design programme.

After considering the range of test materials suitable for use at this level, the AH4 Group Test of General Intelligence was selected. Although appearing somewhat old-fashioned and probably culturally biased, for the purposes of this study it had the advantages of being reputable, relatively inexpensive, and easy to administer. In addition, students on GNVQ Art and Design programmes completed the Spatial Ability test from the NFER Nelson General Ability Tests series.

Basic Skills Test

All colleges were already using the *Assessing Reading and Maths* screening test published by ALBSU, and in some colleges all new students were being asked to take this test as a matter of college policy.

Although this test is designed for use with students whose level of skills is significantly lower than would be expected for entry on GNVQ Advanced programmes, it was agreed that the advantages of familiarity to staff and ease of administration outweighed the disadvantages of a possibly inappropriate level.

The relatively high proportion of students for whom scores were obtained and the variation shown in the scores vindicated this decision.

Socio-economic data

The primary source for the collection of socio-economic data was a questionnaire completed by individual students. In addition to items covering parents' education and academic attainments (or level of education), the questionnaire included other areas of interest to the study, such as attitudes, motivation and aspirations.

In the course of planning the project, the attention of researchers had been drawn to the use by some colleges of postcodes as an easy way of identifying groups of students who appeared regularly to do better or worse than would have been expected from their GCSE scores. A possible explanation for this was that the postcodes indicated the socio-economic characteristics of the areas in which the students were living, and that these were factors affecting the students' achievements.

The project offered an opportunity to test this explanation by matching postcode information to socio-economic data derived from the national census. (Postcodes can be matched more or less accurately to census Enumeration Districts.)

The information on students' socio-economic background derived from items included in the questionnaire used in the project was compared with the socio-economic status based on their postcodes. Socio-economic data based on postcodes were also considered independently as input data in the final analysis of value added.

Course tutors' predictions

The tutors of courses included in the study were also asked at the beginning of the course to provide their own predictions of the grades which they expected their students to achieve at the end of the course. The reason for making this request was the possibility that through the admissions process tutors might have arrived at a more 'holistic' assessment of their students' capabilities than would be reflected in bare examination and test results. Tutors might, for example, have formed impressions of students' motivation at interview, or have been aware of additional information provided in students' records of achievement or portfolios. Tutors were asked to provide a prediction for each student on a seven-point scale, covering the range from distinction to failing to succeed. Feedback on this item showed that many course tutors did not at the beginning of the course think in terms of the final grades they expected for their students. A number were reluctant to provide a prediction until they had got to know their students better. The validity of the data provided in response to this request was therefore questionable in terms of the measurement of value added.

The collection of input data

Collection of input data for the project proved far more difficult than had been anticipated. The study was planned on the assumption that the starting point would be the Individualised Student Records (ISRs) created on the colleges' management information systems (MIS). However, 1994–5 was the first year in which student data was collected through the ISRs, and most college information managers were unable to cope with both satisfying the FEFC's requirements and providing data for the value-added project. Only one college made effective use of its ISRs to provide information for the project.

The main source for the collection of additional personal information was the questionnaire, which was completed by 62% of the students in the study. The information thus obtained was supplemented by an additional form specifically designed to obtain details of students' qualifications on entry.

Characteristics of the student cohort

When the exercise of collecting input data had been completed, the information was analysed to provide feedback on the characteristics of the student cohort for the project managers and the participating colleges. The main findings were:

- the average GCSE grade of students included in the study was between C and D. The average GCSE grade for English, English Literature, and Mathematics was C
- students' scores in the Basic Skills tests and the Psychometric tests were closer to each other than either of them was to their GCSE scores
- commonsense appeared to be reflected in the data, e.g. GNVQ Business students said in the questionnaire that they hoped to work in Finance and Banking
- socio-economic data derived from the questionnaire in general did not correspond well with that derived from postcodes/census information.

The collection of output data

From the beginning of the planning of the study, it had been recognised that the very limited range of outcome measures provided for GNVQ programmes – Pass, Merit, or Distinction – would make it difficult to provide statistically meaningful measurements of value added. With only three grade-based scores available, nearly all students with a pass grade would appear to have underperformed, and nearly all students with a distinction grade would appear to have overperformed.

With the advice of the project advisory committee and in consultation with staff from the GNVQ awarding bodies, project staff explored the possibilities of overcoming this difficulty. The use of informal college unit-based assessments and unit test scores as more discriminating measures had to be rejected as highly problematic.

The additional output data which remained to be considered were the number of GNVQ units passed by students who failed to complete the qualification within two years, and the additional qualifications, such as additional GNVQ units, NVQ units, and GCE AS and A-levels passed by successful students.

It was also agreed that, though not necessarily meaningful in terms of value added, retention data should be collected with the aim of finding out to what extent student retention could be predicted from the input data.

The design for the collection of output data was finalised when the project staff met college representatives in the summer of 1996. Print outs were sent to the colleges listing the students included in the project and making provision for the collection of the following data:

- census date attendance: four census dates were included, from February 1995 to May 1996
- GNVQ grade awarded
- number of GNVQ units passed: mandatory + optional
- number of additional GNVQ units passed
- number of NVQ units passed
- number of additional GCSE subjects passed
- grade of A/S levels passed
- grade of GCE A-levels passed.

The college staff responsible experienced almost as much difficulty over collecting the output data as they had over collecting the input data. The quality of the attendance data was so poor that hardly any use could be made of it. A significant number of staff in colleges appeared to misunderstand the meaning of a 'unit' in the GNVQ. And the information provided on additional qualifications was clearly incomplete. (The most surprising feature of the exercise was the submission of 18 additional students' records, which had been mislaid two years previously!)

In these circumstances, the information provided required careful interpretation. The student sample (now numbering 1970) was divided into the following four categories:

- No record: students for whom there was no evidence in the output data returns that they should have ever been included in the sample – **308**.
- Withdrew: students for whom there was evidence that they started the course, but then withdrew – **647**.
- Not yet completed: students recorded as attending in May 1996, but not achieving a GNVQ pass grade – **240**.
- Completed: students to whom a GNVQ qualification was awarded – **775**.

Analysis of student retention

The characteristics of students who had completed the course were compared with those of students who had withdrawn. Only two of the input items, both from the student questionnaire, significantly predicted course completion:

- determination to complete course (very determined / nothing will stop me)
- importance of course to student (it's very important / it's essential to my plans)

Only a few students who had completed these items on the questionnaire had indicated a weak commitment to completing their courses, whereas more than half withdrew or failed to complete. This finding is therefore not very useful.

The reasons students give for withdrawing from courses vary widely. There is some evidence from other studies that retention rates are influenced by processes which occur during the delivery of the course. This study provided little evidence that retention can be predicted from evidence available at the start of the course.

Data analysis

The statistical analysis of the data was based on the following scoring system:

Input Scores:
GCSE: Grade A = 7 to Grade G = 1

Output Scores:
GNVQ Advanced Grades:
Pass = 4 Merit = 8 Distinction = 12
Additional Level 3 (G)NVQ Units: 1/3 of a point
Additional GCSE Passes (A–C): 1/3 of a point
Additional Level 2 NVQ Units: 1/6 of a point
GCE A-level Passes: UCAS points score (A = 10, B = 8, etc)
GCE A/S-level Passes: UCAS points score (A = 5, etc).

For students who did not complete the GNVQ qualification, points given for each completed unit up to a maximum of 12: 1/3 of a point.

The study provided potentially a very large number of variables for analysis. Priority was given to examining the relationship between the three main cognitive variables (GCSE scores, psychometric test scores, and basic skills test scores) and the grade scores achieved by those students who completed the qualification. All the evidence from earlier studies of this kind indicated that the outcomes would be primarily determined by previous attainments and/or ability. The significance of additional factors (using multiple regression techniques) would be apparent only if the initial analysis provided a basis for more refined discrimination.

Key findings

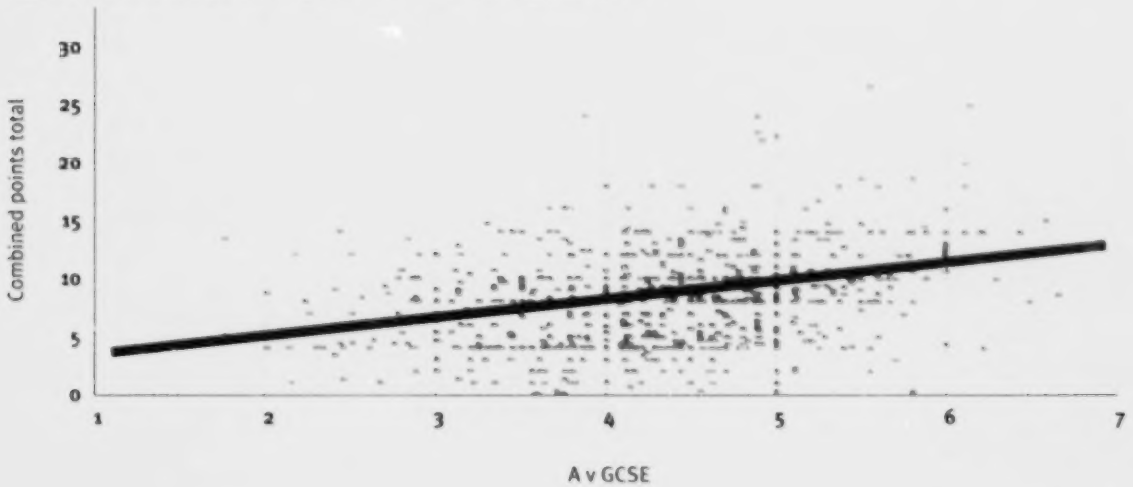
The study did not provide any simple solution to the measurement of value added on vocational courses. Neither alternative nor additional data appeared to contribute an improvement on the low level of correlation found in previous studies using GCSE scores as input data.

The most interesting and useful findings to emerge from the analysis of a relatively complex set of data were:

- The best predictor of GNVQ attainment on all four programmes was average GCSE score. The correlation was of the order of $r = 0.3$, indicating that GCSE scores would on average account for only 9% of the variation in GNVQ attainment.
- Analysis of the value added for a selected group of larger cohorts of students following specific GNVQ programmes in individual colleges showed correlations ranging between $r = 0.15$ and $r = 0.52$. The average correlation between GCSE and GNVQ scores for these selected larger cohorts appeared to be significantly higher than the overall correlations obtained for the study as a whole. The wide differences between correlations found for different groups in different colleges may indicate an area for further investigation.
- On the GNVQ Art & Design programme, spatial ability test scores appeared to provide as good a prediction of success as GCSE scores.
- Previous success on a GNVQ Intermediate or BTEC First course appeared to predict success in GNVQ Advanced (although the numbers were very small).
- Course Tutors' predictions were also statistically significant, but no better at predicting outcomes than GCSE scores.
- Postcodes could not be used as a substitute for socio-economic data obtained directly from individual students.

Figure 3 shows in graphical form the results of the analysis using the scoring system described above.

Figure 3 Plot of average GCSE against total score with regression line



Technical conclusions

Analysis comparing the input data from GCSE scores, psychometric tests, basic skills tests, and course tutors' predictions with GNVQ attainments showed that most of the inputs were statistically significant (with a p-value of less than 0.00) in the expected direction, but not significant enough to allow useful value-added measurement to be carried out.

More detailed analysis using a wider range of data produced findings which were difficult to interpret. Some of these findings appeared to be spurious.

The failure in this study to find a method of measuring value added for GNVQs seems to require some explanation.

In the course of collecting data for this study, there was some evidence that different colleges were offering a diversity of programmes under the umbrella of the same formal qualification. In one case, the GNVQ qualification appeared to be a subsidiary target for students whose primary aim was to achieve qualifications at GCE A-level. In another college students achieved an exceptionally high proportion of distinction grades although they had entered with poor GCSE grades and low expectations on the part of their course tutors.

Adult A-level value-added feasibility study

Seven of the ten colleges which took part in the value-added feasibility study for GNVQ Advanced programmes also participated in an investigation of value added for students over 19 enrolling in September 1994 on GCE one-year A-level courses.

The courses included in this study were: English, Human Biology, Law, Mathematics, Psychology, and Sociology. The size of the sample was 650 students, of whom the majority (511) were on courses at Bournemouth and Poole College of Further Education.

The main findings of this study were:

- The ALBSU test score was the strongest predictor of A-level success, with a correlation of $r = 0.33$ based on a sample of 239 students.
- When the correlations were analysed separately for each of the six A-level subjects, the ALBSU test score was the strongest predictor for only three of them (Law, Mathematics, and Sociology); AH_4 test scores were the strongest predictor for Human Biology and Psychology, and average GCSE scores for English.
- With this sample other variables could not be used to strengthen the predictions of A-level grade provided by ALBSU test scores.
- As with the GNVQ Advanced cohort, the students who had expressed a strong commitment at the beginning of their course were more likely to complete it.
- Students who were taking the course for the second time were more likely both to complete it and to attain higher grades.

The way forward

For national bodies

Further research using national databases

The difficulties of data collection and the lack of positive findings from this study indicate that the methodology adopted does not provide a cost-effective way of measuring value added.

The finding that GCSE results provide the best available prediction of achievements on GNVQ Advanced courses, albeit a relatively poor one, may be worth following up. National databases of students on GNVQ courses are now being developed, and the information these contain can be linked to national information on GCSE results. In this way it would be possible to trace the development of the relationship between GCSE and GNVQ results. The fact that the system of delivering and grading GNVQ Advanced qualifications is currently under review may make this an attractive course of action, once the changes have been introduced, since the schemes of assessment which emerge may be more susceptible to effective value-added measurement.

For colleges

Local investigations of value added

Some colleges, which have carried out their own investigations of the measurement of value added, have found that factors specific to the area and community within which they are working may be used to predict performance and identify need for support. The interpretation of such findings is often difficult. But, in the absence of a universal methodology, colleges should not be discouraged from using such local value-added measurements, provided that they are based on a sound statistical methodology.

Awareness of other research findings

Indications of areas which may be worth investigating are to be found in reports on value-added research in educational journals and the press, and in findings published by the DfEE and other organisations which analyse data on a large scale. Findings of research carried out in schools as well as colleges may both help to identify local factors which may be significant and assist in the interpretation of findings which seem to be problematic. Although many people find the technicalities of such research difficult to follow, the findings are usually quite simple to grasp, and may be valuable in drawing attention to factors which influence retention and achievement.

Using target minimum grades to improve retention and achievement

The development of GCE A-level tutorial systems, which focus on the achievement of targets based on value-added predictions may provide a useful model for the provision of support on other programmes. Target grades do not necessarily have to be derived from statistical measurement in order to motivate students and monitor performance. What is important is that the agreed target grade should be realistic in terms of the student's past performance and abilities, and should be sufficiently precise to be monitored against actual achievements.

Even on programmes for which statistically based predictions are available, they may not provide appropriate targets for every student. For example, a student enrolled for an A-level course may, for reasons such as illness, have seriously underperformed in GCSE examinations; in such a case, the student's target should be higher than that based on statistical prediction. This case is no different in principle from that of courses for which value-added measurement does not provide clear indications of appropriate target grades. The real strength of this approach lies in a clearly focused, well-managed tutorial support system.

References

Adult Literacy and Basic Skills Unit (1993) **Assessing reading and maths: a screening test.** ALBSU

A-Level Information System, Curriculum Education and Management Centre, University of Durham

Audit Commission/HMI (1992) **Can value-added indicators be applied to vocational courses?** Audit Commission

Audit Commission/HMI (1993) **Unfinished business.** HMSO

DfE (1995) **Value added in education – a briefing paper.**

DfEE (1997) Statistical Bulletin Issue no 1/97 **GCSE and GCE A/AS level performance of candidates attempting two or more GCE A-levels or AS equivalents in 1995/6.** Stationery Office

DfEE (1998) Statistical Bulletin Issue no 3/98

FEU (1993) **Value added in further education.**

FEU (1994) **Value added – an update.**

FEU/FEDA (1995) **Current developments in value added.**

Heim A W, K P Watts and V Simmonds (1981) **AH4: group test of general intelligence.** NFER-NELSON

Smith P, C Whetton (1988) **General ability tests: spatial ability.** NFER-NELSON

Further reading

Colleagues may also find the following publications useful:

Foster P, V Howard and A Reisenberger (1997) **A sense of achievement – outcomes of adult learning.** FEDA

Green M and D Bartram (1998) **Initial assessment to identify learning needs.** FEDA

Martinez P (1998) **Student retention: case studies of strategies that work.** FEDA

Acknowledgements

The project work on which this report is based was managed by Stella Dixon, and the project consultant was Philip Barnard.

Technical support for the statistical analysis of data was provided by the Curriculum Evaluation and Management Centre (CEMC), now based at the University of Durham, under the direction of Professor Carol Fitzgibbon, working in collaboration with Dougal Hutchison, Head of Statistics, NFER.

FEDA is grateful to the staff and students of the following colleges which took part in the feasibility study:

- Blackpool and the Fylde College
- Bournemouth & Poole College of FE
- Burton upon Trent Technical College
- Gateshead College
- Llandrillo College
- Milton Keynes College
- South Tyneside College
- Walsall College of Arts & Technology
- West Notts College
- Wirral Metropolitan College

FEDA is also grateful for the support given by members of the project Advisory Group:

- Ms Gill Brain
- Mr Dave Sambrook
- Ms Helena Stockford
- Mr Bob Challis
- Mr Robert Ensum
- Ms Caroline Kempner
- Ms Elaine Carabok
- Ms Kate Anderson
- Dr Sally Thomas
- Mr James Kennedy
- Mr John Gardner
- Ms Audrey Brown
- Ms Janet Cormack
- Ms Lesley Virgin
- Ms Joyce Johnston
- Mr Peter Goff
- Ms Louise Stern

Published by the **Further Education Development Agency (FEDA), Dumbarton House, 68 Oxford Street, London W1N 0DA**
Feedback and orders should be directed to **FEDA Information Centre, Citadel Place, Tinworth Street, London SE11 5EH**
Tel: 0171 962 1280 Fax: 0717 962 1266 Registered with the Charity Commissioners
Editor **Jenny Rhys** Designer **Dave Shaw** Printed by **Blackmore Ltd, Shaftesbury, Dorset**